

# **Use of Statistics in Evaluation of Trace Evidence**

**Robert D. Koons  
CFSRU, FBI Academy  
Quantico, VA 22135**

**Trace Evidence Symposium**

**8/15/2007, Clearwater Beach, Florida**

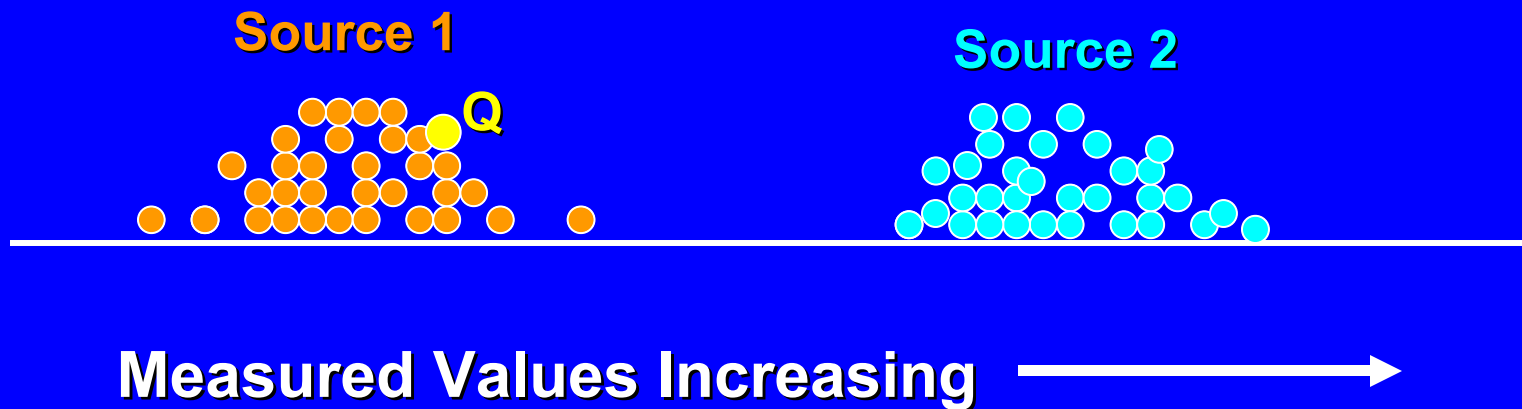
# My rules for comparison of trace evidence

- Comparison of trace evidence is best thought of as a process of elimination.
- Selection of features for comparison that provide the best source discrimination is *always* a good idea.
- Match criteria do not have to be statistically-based to be effective.
- Frequency of occurrence statistics for trace evidence can *almost* never be calculated for good discriminating features.
- Databases are useful for making broad classification rules, but they are generally useless for calculating the significance of a match.

# Characteristics of Measurements on Trace Evidence

- **Data for many variables are “continuous”**
- **Data distributions are “often” unknown**
  - **Frequency distributions are nonstandard**
  - **Across-sample distributions are unknown**
  - **Accuracy and precision of data depends on analytical method**
  - **Databases are both time and location dependent**
- **Forensic and scientific (statistical) issues may not be the same**

# Significance of a Match



# Fisher's Ratio

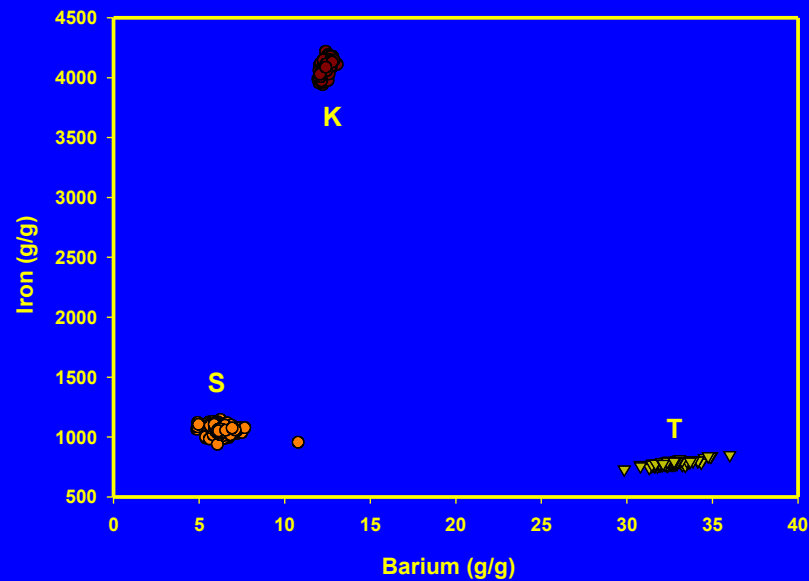
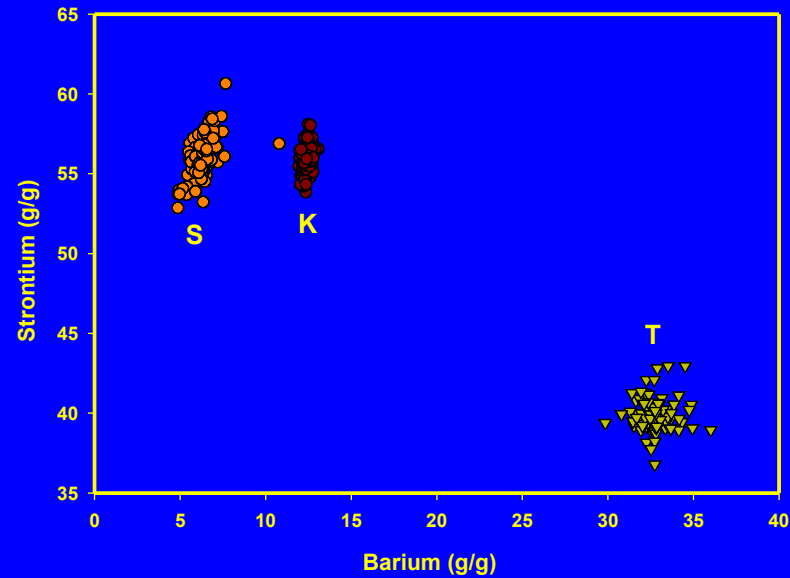
measure for linear discriminating power of a variable

$$F = \frac{(m_1 - m_2)^2}{v_1 + v_2}$$

$m_1$  and  $m_2$  are the means of class 1 and class 2

$v_1$  and  $v_2$  are the variances

# Three sheets of float glass

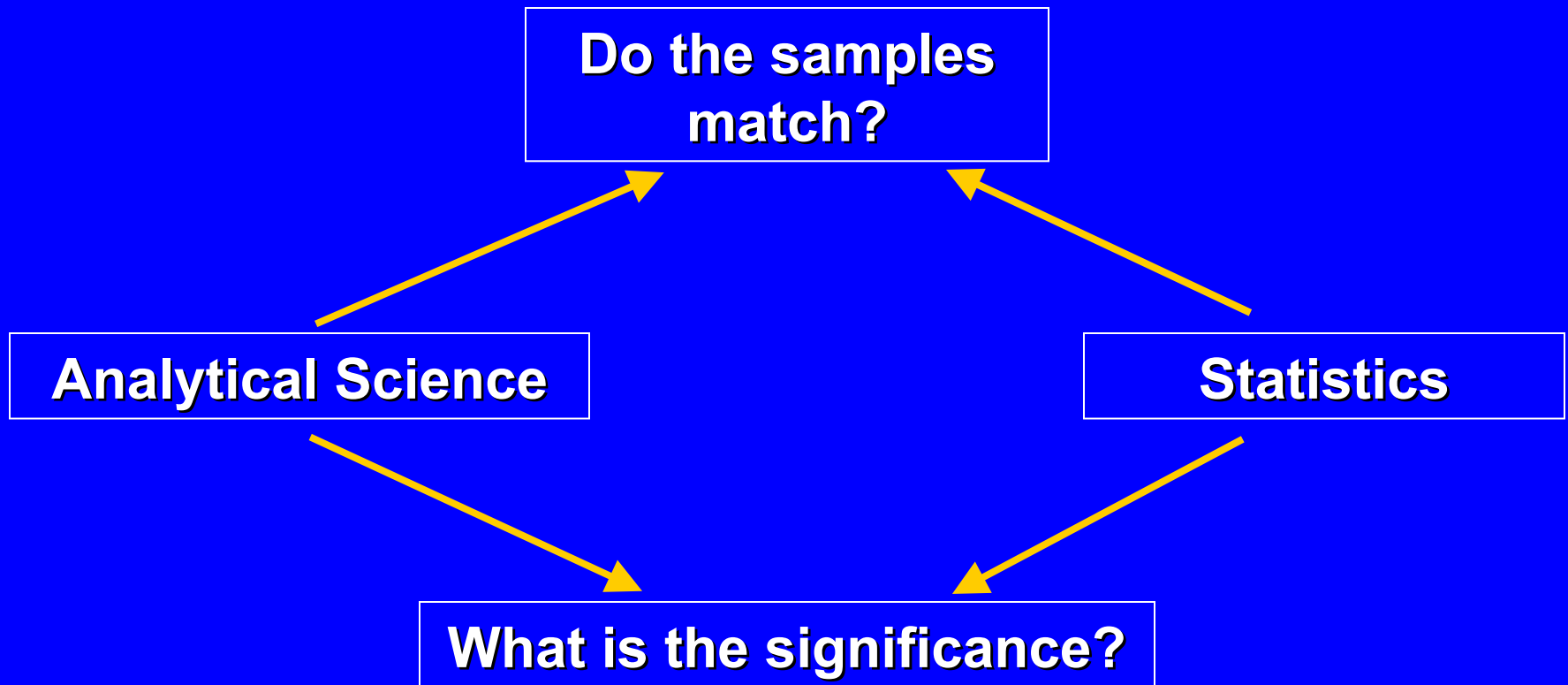


# Truth Table

Truth

Decision	Truth	
	Same Source	Different Source
Indistinguishable	Correct	False inclusion Type II error
Exclusion	False exclusion Type I error	Correct

# Roles in Sample Comparison

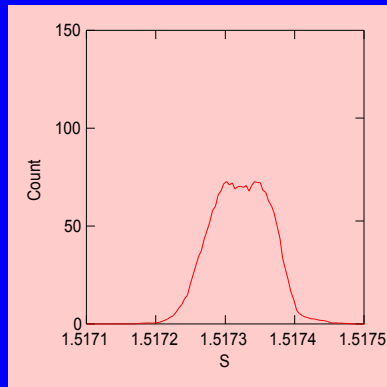




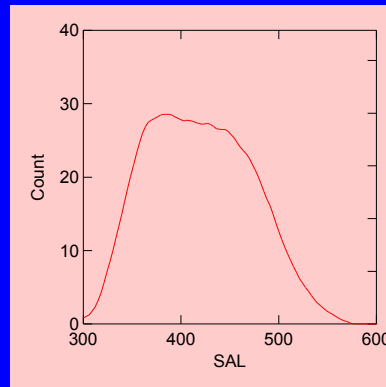
# Some Match Methodologies

- ***t*-test**
  - Welch's modification?
  - Multivariate (Bonferroni) correction?
- **Range overlap (many to many or one to many)**
- **$2\sigma$ ,  $3\sigma$ , etc. (or  $2s$ ,  $3s$ , etc.)**
- **Continuous probabilistic approach**
- **Dimension reduction, then match**
- **Cluster analysis**
- **Multivariate test (Hotelling's  $t^2$ )**
- **Discriminant analysis (PCA)**

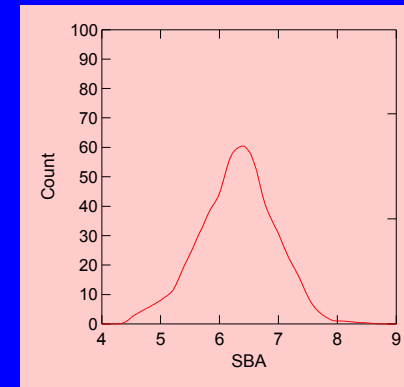
# Measured Distributions in a Sheet of Float Glass



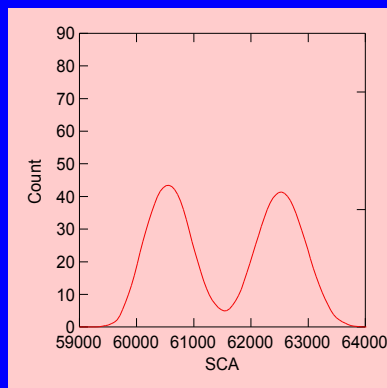
**Refractive Index**



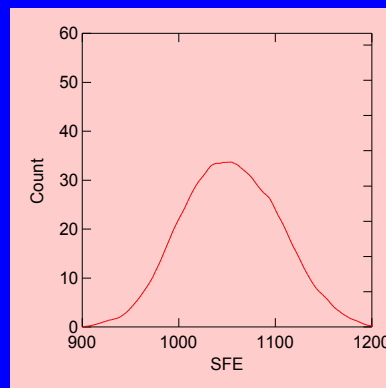
**Aluminum**



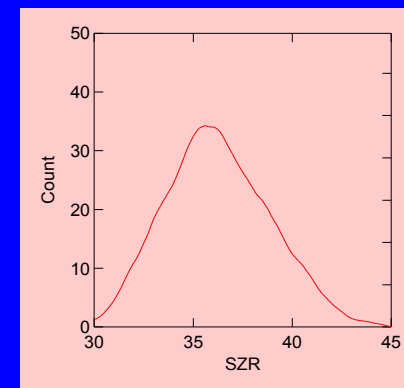
**Barium**



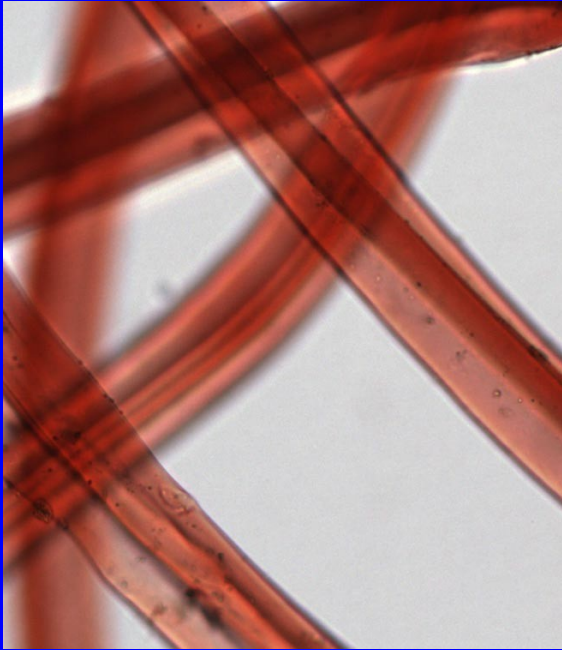
**Calcium**



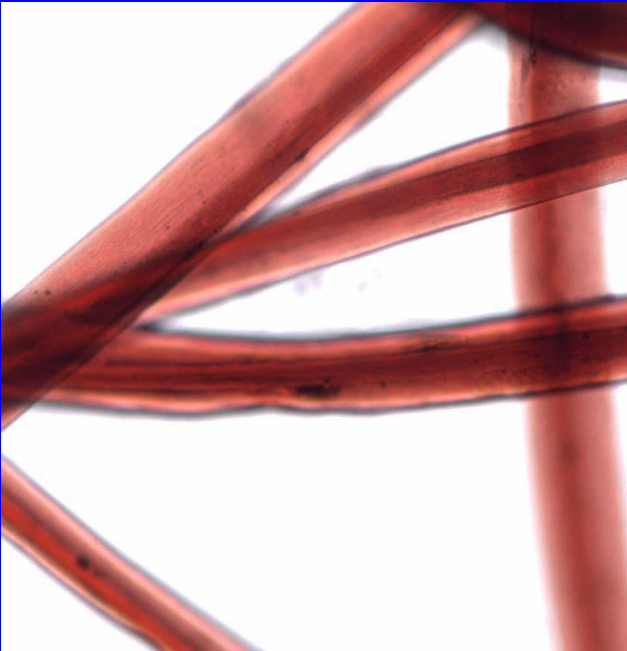
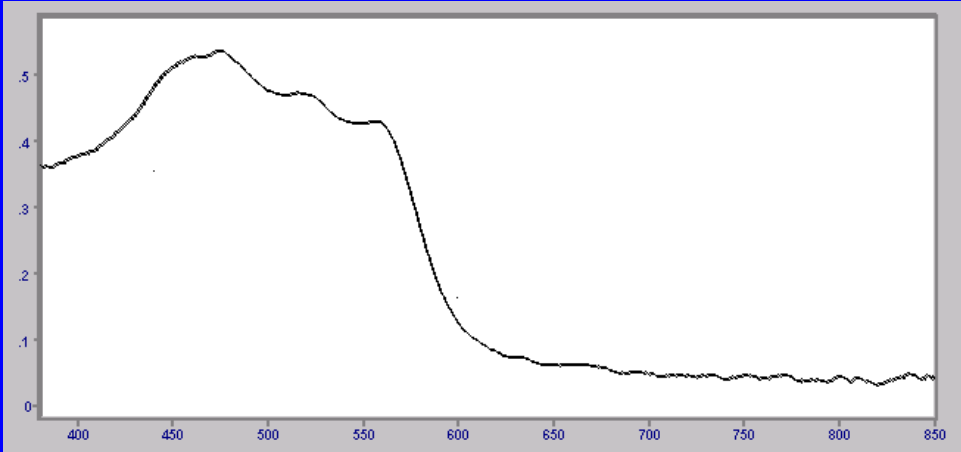
**Iron**



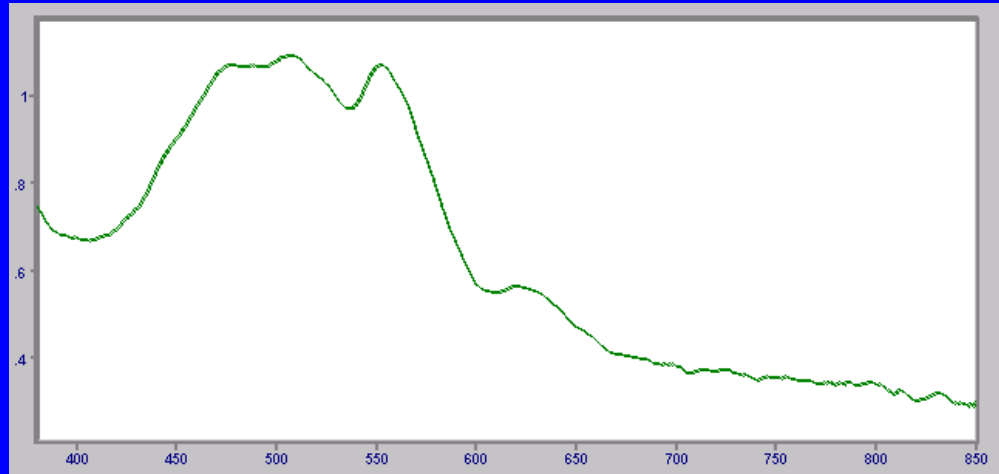
**Zinc**

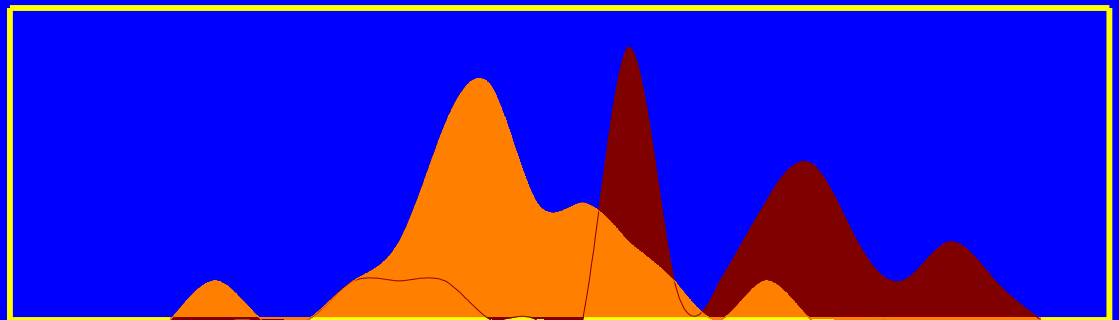
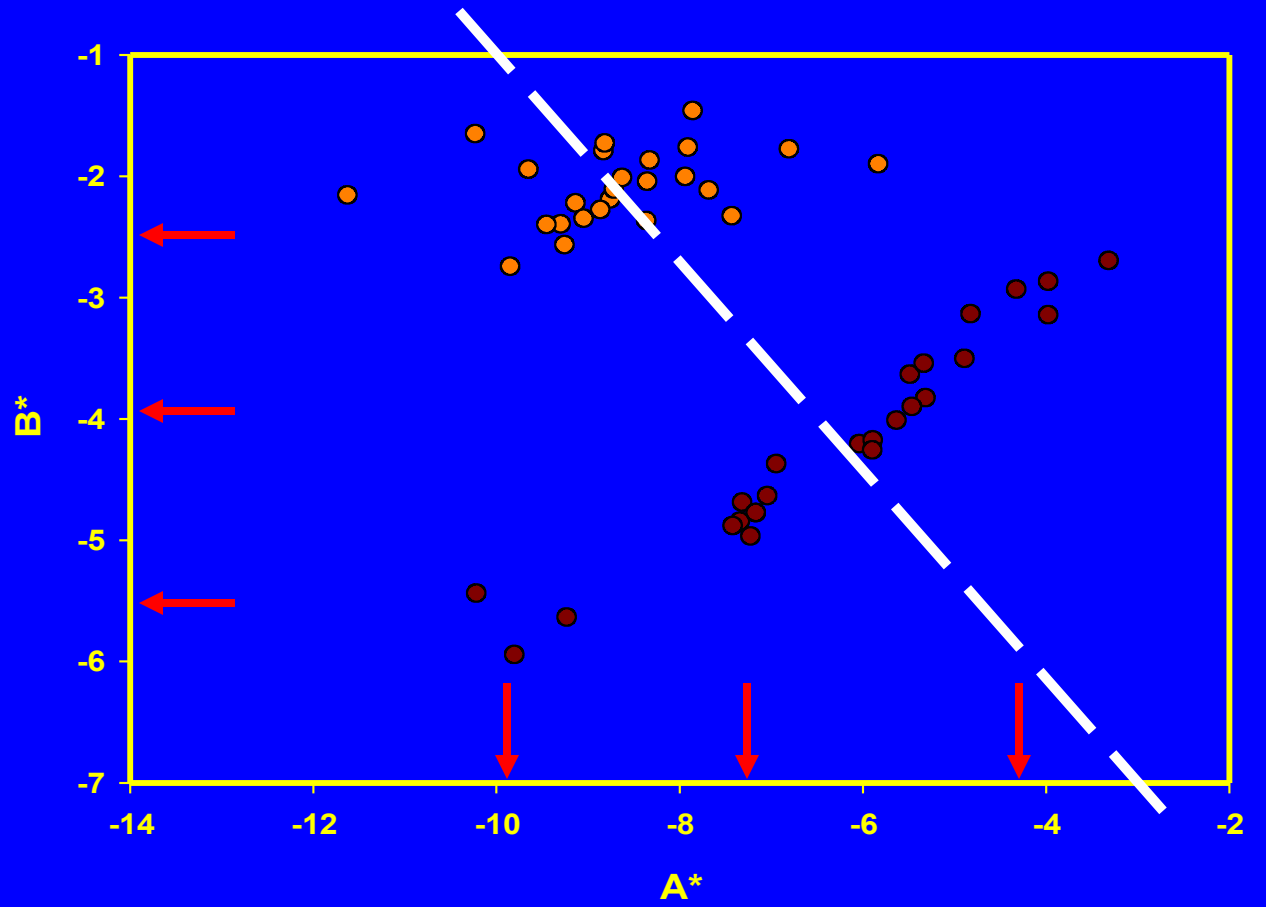
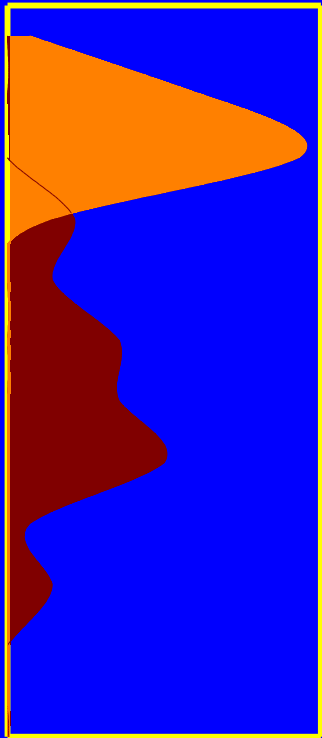


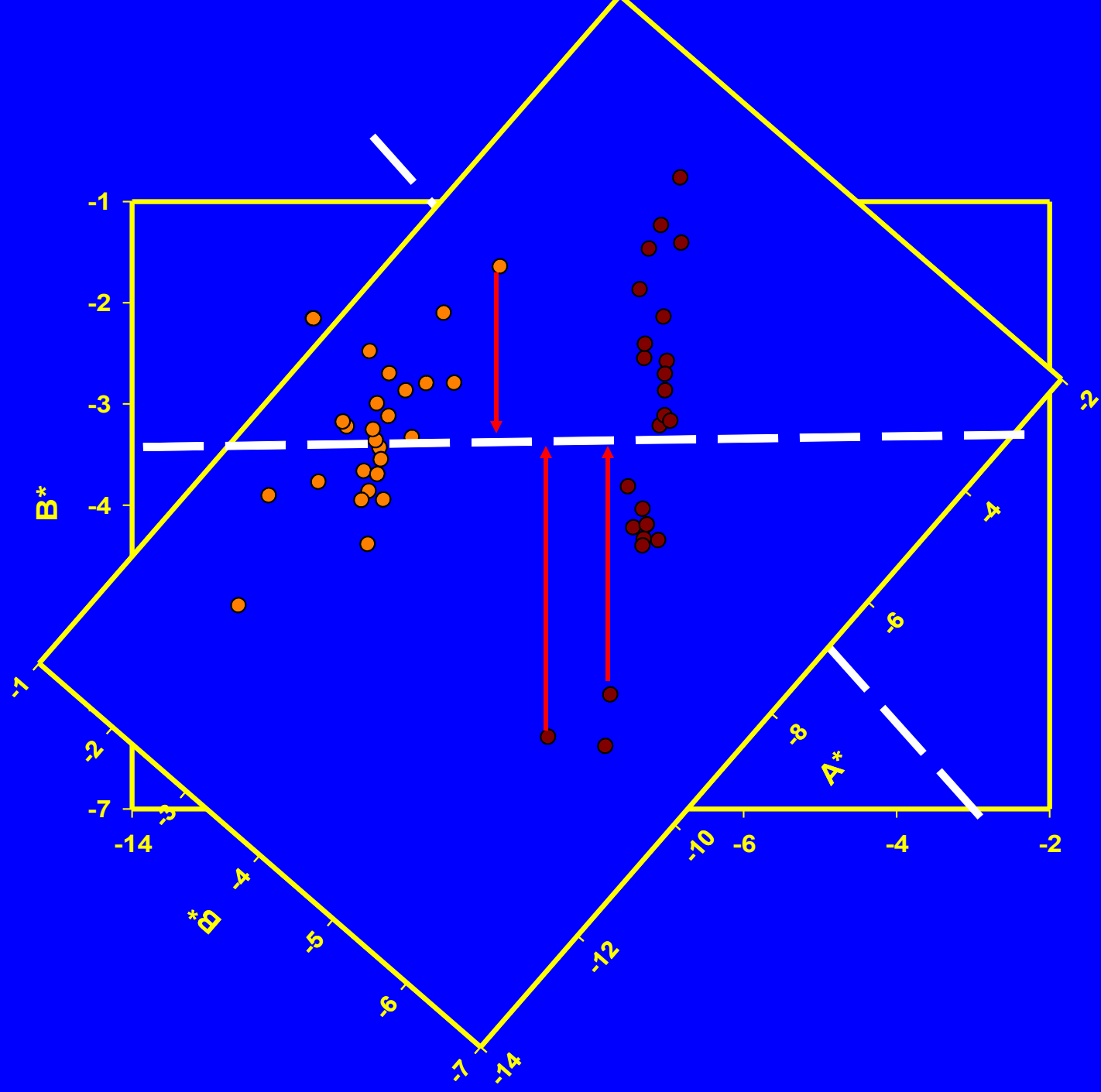
**Fiber No 42**



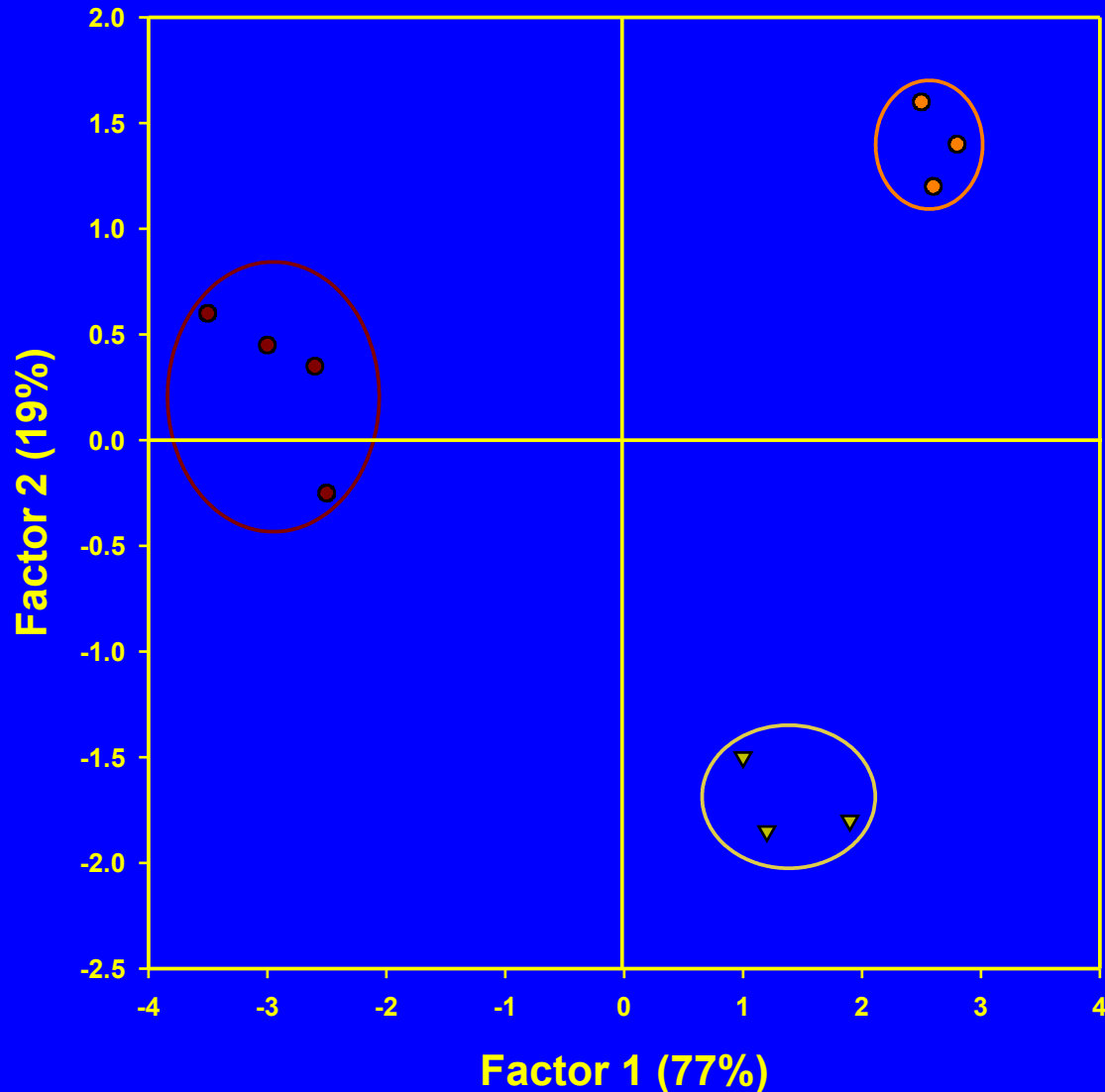
**Fiber No 52**







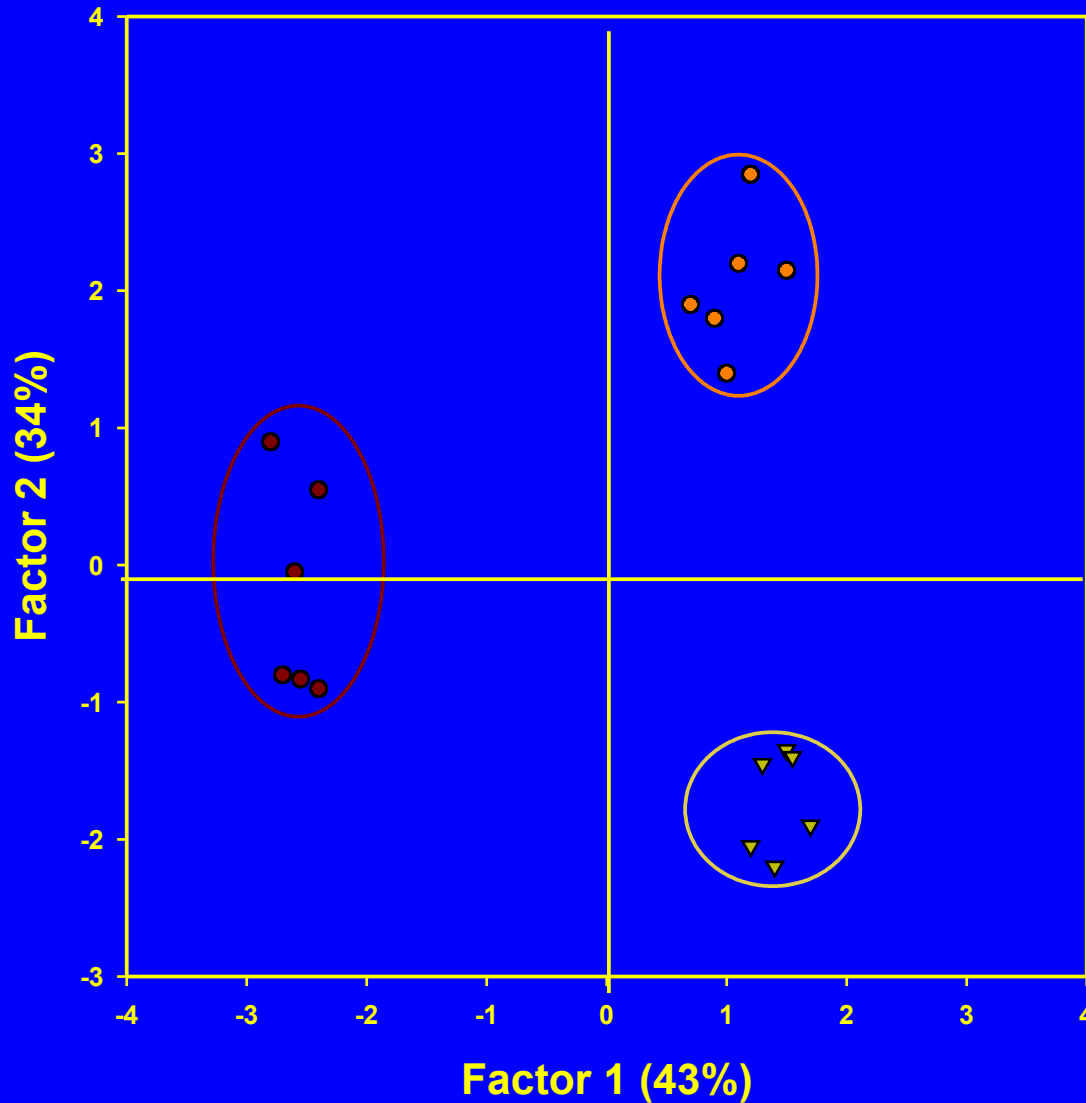
# PCA plot of Australian ocher data



**B, Sc, Se, Rb, Pd,  
Hf, Th, and U in  
ochers from 3 areas  
of Australia**

**From: R.L. Green &  
R.J. Watling, JFS 7/07**

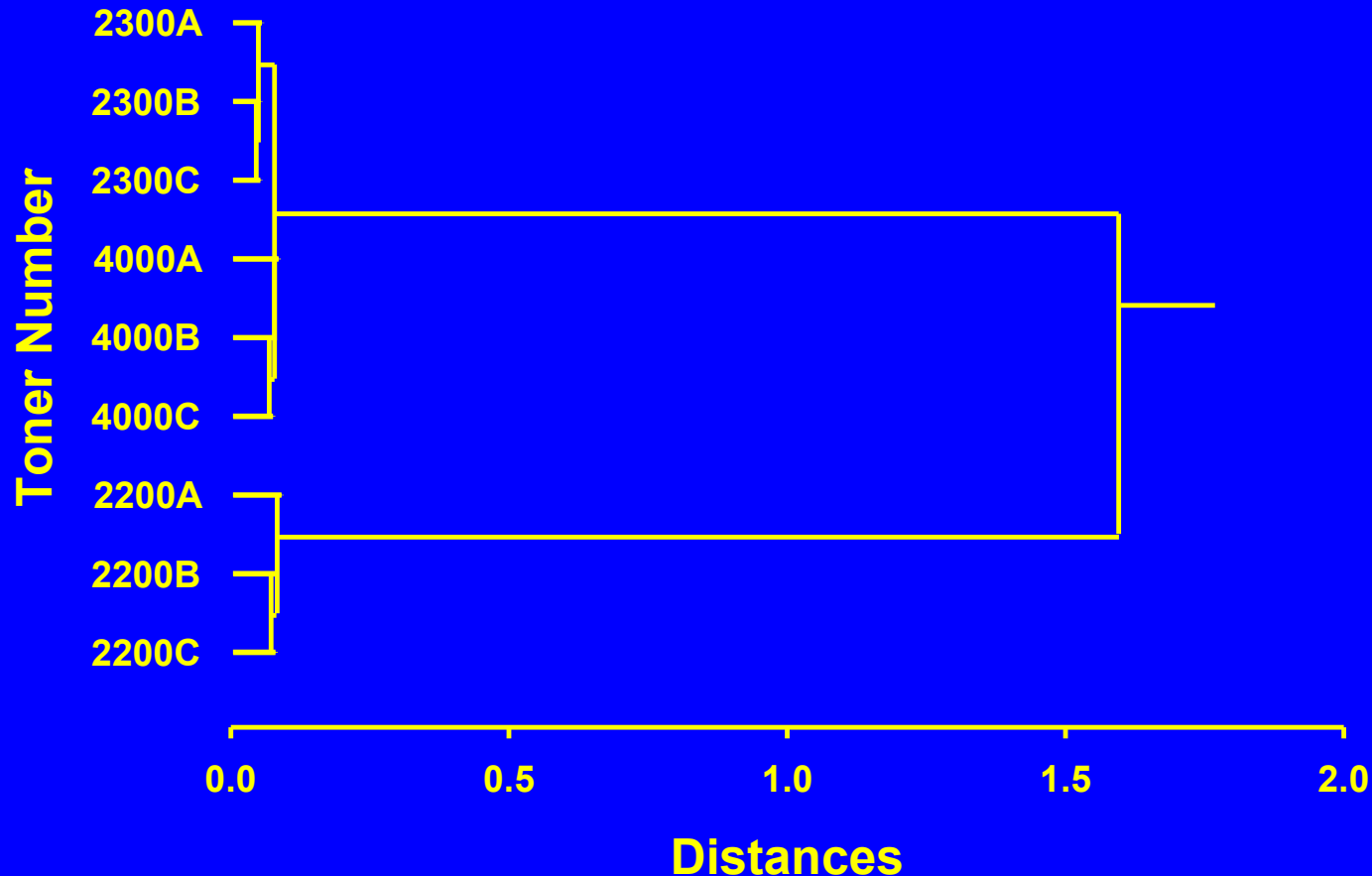
# PCA plot of Australian ocher data



ochers from 3  
populations within a  
single region of  
Australia

From: R.L. Green &  
R.J. Watling, JFS 7/07

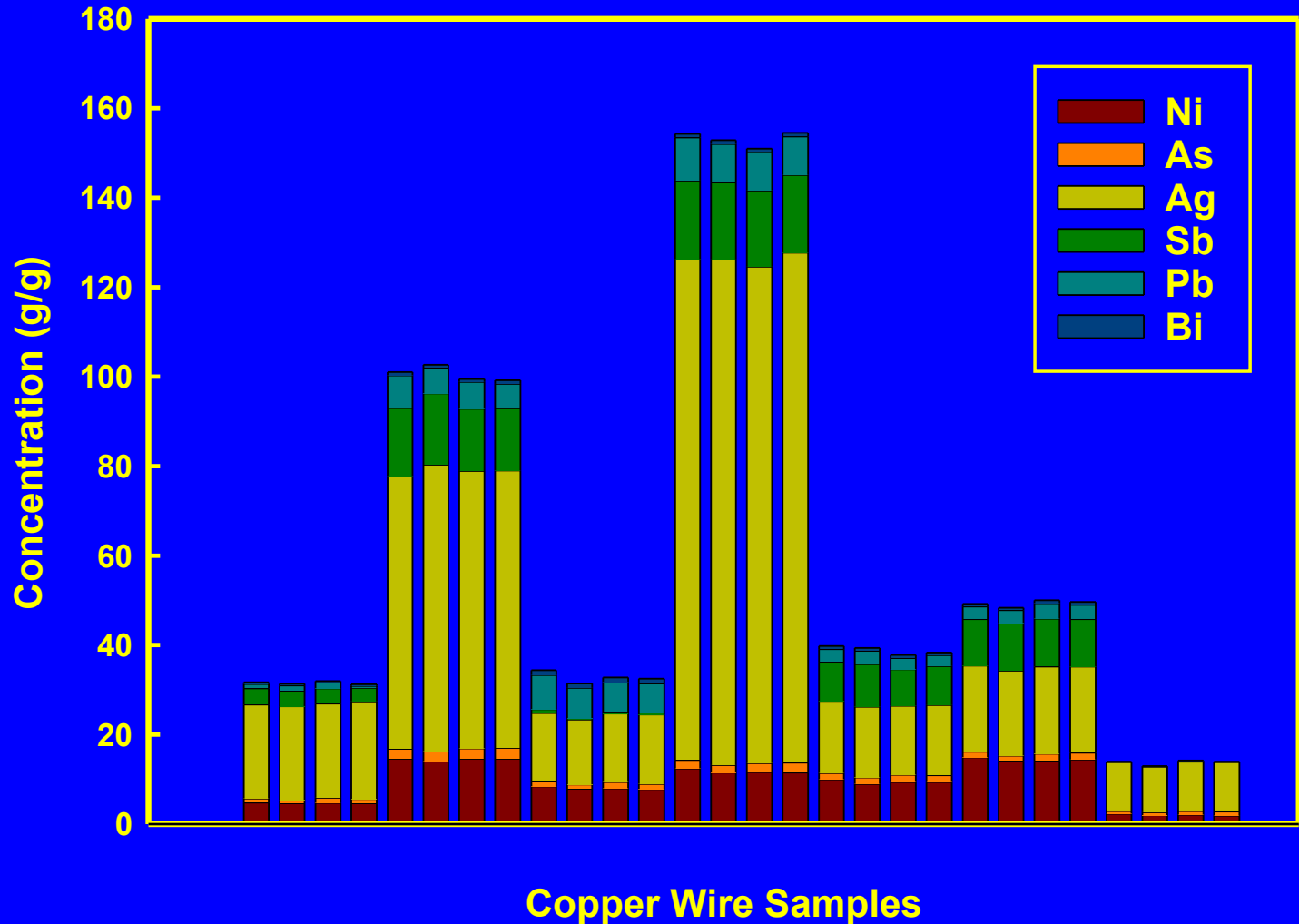
# Classification of laser jet toners



LA-ICP-MS data, heirarchichal clustering, Euclidean distance, 9 elements, normalized



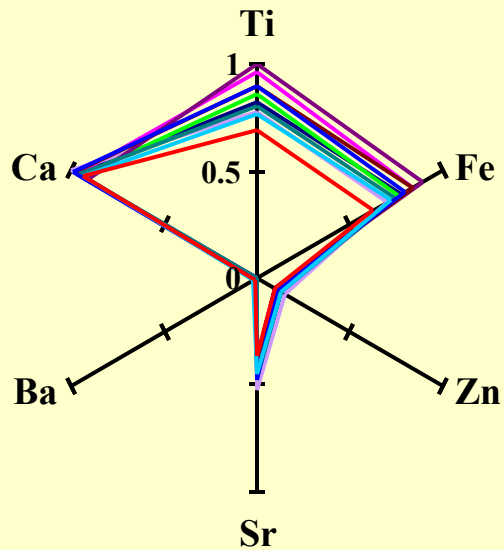
# Seven-Stranded Copper Wire by ICP-MS



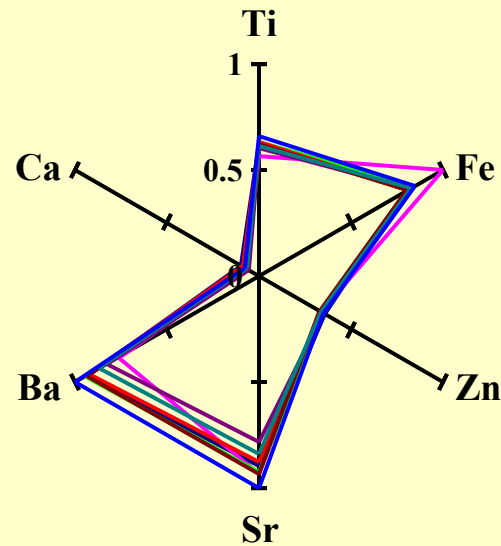
# Star Plots

## Polyethylene Trash Bags

### Hefty Easy Flaps



### Glad Quick Tie



### Ruffies

